La documentación lingüística en la investigación de lenguas en la Amazonía peruana

Language Documentation in Linguistic Research in the Peruvian Amazon

Jairo Valqui
Universidad Nacional Mayor de San Marcos
jvalquic@unmsm.edu.pe
Rosario Cosar
Universidad Nacional Mayor de San Marcos
rosariocosar@gmail.com
Pamela Jimenez
Universidad Nacional Mayor de San Marcos
pamelajimenez01@gmail.com
Licett Ramos
Universidad Nacional Mayor de San Marcos
licettseis09@gmail.com
Pier Suclupe
Universidad Nacional Mayor de San Marcos
suc1093@gmail.com

Resumen

Este artículo aborda una experiencia en documentación de lenguas en la selva central del Perú con el propósito de reflexionar sobre algunas dimensiones metodológicas de la documentación lingüística que pueden extenderse a toda la disciplina mayor. Esta experiencia muestra que las metodologías de recojo de datos y el uso de programas informáticos pueden mejorar una serie de análisis lingüísticos específicos dentro de cualquier campo de la lingüista.

Palabras clave: *documentación lingüística, lenguas peruanas*

Abstract

This paper discusses an experience in language documentation in the Peruvian Amazon in order to reflect on some methodological

dimensions that may extend to the linguistic research. The aim of this paper is to show that methodologies and the use of computer programs in language documentation can improve the analysis in any field of linguistics.

Keywords: *Language Documentation, Linguistic Research, Peruvian Languages*

Introducción

La línea de trabajo más extendida en la documentación de lenguas es la que se relaciona directamente con la actividad de registro de prácticas comunicativas de una comunidad lingüística en particular. Esto se debe a que su origen está vinculado estrechamente con iniciativas por registrar y almacenar datos de lenguas que se encuentran en algún nivel de amenaza o en serio peligro de extinción. Es más, reconocidos especialistas en este campo sostienen que estos procesos de documentación pueden ser considerados como una respuesta lingüística importante ante el aumento dramático de lenguas amenazadas en el mundo (Cf. Himmelmann, 2007). En la actualidad, en este fructífero campo, existen repositorios digitales que almacenan muestras de lenguas amenazadas de diferentes partes del mundo.

La otra línea de trabajo en la documentación de lenguas es la que se vincula con una actividad tradicional en la investigación lingüística: la descripción. Para Dimmendaal (2010), la lingüística descriptiva recibió un nuevo impulso como resultado del trabajo con lenguas amenazadas a fines de la década de los 80 y con la incorporación de tecnologías de la información durante la década de los 90. Dentro de esta otra rama de la documentación lingüística, no solo se resalta el papel de la descripción al momento de procesar los archivos digitales de audio y video con anotaciones lingüísticas, sino más cuando estos datos ya procesados sirven como soporte para los análisis en diferentes aspectos del campo disciplinar de la lingüística (Cf. Himmelmann, 2012).

Este artículo aborda una experiencia en documentación de lenguas en la selva central del Perú con el propósito de reflexionar sobre algunas dimensiones metodológicas de la documentación que pueden extenderse a toda la disciplina lingüística en general. Esta experiencia muestra que las metodologías de recojo de datos y el uso de herramientas modernas pueden mejorar una serie de

análisis en diferentes niveles de la investigación lingüística. El documento está organizado en tres secciones: en §1 se presenta un balance de los proyectos de documentación lingüística en el Perú, en §2 se detalla la metodología de recojo de los datos de esta experiencia en particular, y en §3 se muestran los resultados de la documentación y se discuten sobre sus implicancias en el campo de la lingüística.

1. Antecedentes

Desde sus inicios, la documentación lingüística se ha enfocado en el registro, almacenamiento y accesibilidad de datos de lenguas en peligro de extinción. Su preocupación, en ese sentido, se orientó a resolver cuestiones en torno a los procesos de recojo de datos (Austin, 2007), las metodologías de trabajo en y con las comunidades lingüísticas (Mosel, 2007), el trabajo colaborativo (Glenn, 2009), entre otras dimensiones. Como así lo entiende Himmelmann (2012), la documentación lingüística se ha ocupado principalmente de los datos preliminares, de los datos primarios y de sus interrelaciones que pueden ir desde las maneras de registrar y almacenar los datos hasta los procesos de crear transcripciones y traducciones.

En la actualidad, el desarrollo de este campo puede ser reconocido a través de repositorios internacionales como Elar (Endangered Languages Archive)[1] y Dobes (Dokumentation Bedrohter Sprachen)[2] que se encargan de almacenar datos primarios conformados por archivos en audio y/o video con anotaciones lingüísticas de lenguas en algún grado de amenaza o en peligro de extinción. Muestras de estas lenguas incluidas en estos repositorios alcanzan al Perú en donde vienen documentándose 15 lenguas que representan a 8 familias lingüísticas como el Arawak, Bora, Kandozi, Huitoto, Pano, Shimaco, Tupí-guaraní y Záparo.

Hasta el momento, en ambos repositorios se han almacenado 3126 archivos de audios, 1592 archivos de video y 945 anotaciones como se muestra en el Cuadro 1. En la parte superior de este cuadro, el repositorio Elar consigna 9 proyectos de documentación que están trabajando con 9 lenguas amazónicas: asheninka del Perené, caquinte, cashibo-cacataibo, chapra, mayoruna, nahua, iquito, kokama-kokamilla y urarina. En este repositorio, se han almacenado un total de 953 archivos de audios, 199 archivos de video y 81 anotaciones (en 3 proyectos aún no se presentan datos). En la parte inferior del cuadro, el repositorio Dobes presenta 2 proyectos de documentación que están trabajando con 6 lenguas amazónicas: bora, ocaina, witoto, resígaro, nonuya y cashinahua. En este repositorio,

se han almacenado un total de 2173 archivos de audios, 1393 archivos de video y 864 anotaciones.

Cuadro 1. Proyectos de documentación lingüística en ELAR y DOBES

	ELAR (The Endangered Languages Archive at SOAS, London)					
	Proyecto	Lengua o variedad	Coordinador	Audio	Video	Anotaciones
1	Asheninka Perene (Arawak) 2010 collection, from eastern Peru	Asheninka Perene	Elena Mihas	73	75	0
2	An Audiovisual Corpus of Caquinte (Arawak)	Caquinte	Zachary O'Hagan	0	0	0
3	A Documentation of Cashibo-Cacataibo of San Alejandro (Pano) with a Focus on Information Structure	Cashibo-Cacataibo	Daniel Valle	293	0	48
4	Documentation of Kandozi and Chapra (Candoshi-Shapra) in Loreto, Peru	Chapra	Simon Overall	0	0	0
5	Documentation of the Mayoruna Languages	Mayoruna	David Fleck	285	0	0
6	Documentation of mythology and shamanic songs of the Nahua	Panoan	Conrad Feather	84	19	0
7	Iquito–Castellano and Castellano–Iquito Dictionary	Iquito	Christine Beier	0	0	0
8	The Kukama-Kukamiria Documentation Project	Kokama-Kokamilla	Rosa Vallejos	88	95	33
9	Urarina Language and Culture	Urarina	Knut Olawsky	130	10	0
			subtotal de archivos	953	199	81
			Porcentaje en relación con el total	30%	13%	9%
	DOBES (Documentation of Endangered Languages)					
10	Cashinahua	Cashinahua	Bernard Comrie	205	55	91
11	Pleople of the Center	bora, ocaina, witoto, resígaro y nonuya	Frank Seifart	1968	1338	773
			subtotal de archivos	2173	1393	864
			Porcentaje en relación con el total	70%	78%	91%
Total de archivos				**3126**	**1592**	**945**

Fuente: Elaboración propia a partir de los datos consignados en cada repositorio

En el ámbito nacional, el recientemente creado Archivo Digital de Lenguas Peruanas[3] almacena datos primarios básicamente en audio con anotaciones del aimara, awajún y kakataibo que, sumados a los anteriormente mencionados, aumentan la cifra a 18 lenguas indígenas en proceso de documentación. La gran mayoría de estos procesos de documentación de lenguas indígenas peruanas están orientadas hacia lenguas amazónicas. Los datos del Cuadro 1 nos revelan que el repositorio Dobes es el que más ha almacenado archivos de audio en un 70%, archivos de video en 87% y archivos con anotaciones en 91% con respecto a los datos totales.

Un dato interesante del cuadro es que el procesamiento de archivos con anotaciones lingüísticas es más lento que el procesamiento de los otros. La cantidad de anotaciones siempre es menor en cada repositorio mostrado. Esto posiblemente lleva a que en un proceso de documentación lingüística no se resalte o no se visibilice el papel de la descripción en dicha actividad. Papel, sin embargo, importante en la medida, que como así lo entiende Himmelmann (2012), la descripción provee el *input* tanto para las actividades de transcripción como en las decisiones sobre qué documentar. Más aún cuando los datos primarios son analizados con el propósito de arribar a generalizaciones en el campo de la lingüística en general. La proyección de esta otra manera de ver la documentación de lenguas puede ser provechosa para la disciplina lingüística en su sentido más amplio.

> Documentary linguistics […] is primarily concerned with raw and primary data and their interrelationships, including issues such as the best ways for capturing and archiving raw data, transcription, native speaker translation, etc. Descriptive linguistics […], on the other hand, deals with primary and structural data and their interrelationships, i.e., primarily with the question of how valid descriptive generalizations can be derived from a set of primary data. […] In this view, documentation has the central task of making description accountable and replicable, and is thus of fundamental importance for making linguistics an empirical science. (Himmelmann, 2012, p. 199).

2. Metodología

La experiencia de documentación lingüística en selva central fue coordinada con la Asociación Regional de Pueblos Indígenas de la Selva Central (ARPI-SC) a través de su Programa de Formación de Profesores en Educación Intercultural Bilingüe y financiada por la Universidad Nacional Mayor de San Marcos (UNMSM) a través del proyecto de investigación N.o 140304041. Esta interrelación fue doblemente beneficiosa: por un lado, se contó con la predisposición y compromiso de los sabios y jóvenes por rescatar y preservar aspectos culturales y lingüísticos de los pueblos ashaninka, nomatsigenga y yanesha dentro de este programa de formación de docentes en EIB y, por otro lado, se aprovechó la experticia de los comunicadores sociales y lingüistas para la obtención y procesamiento de datos en los procesos de documentación de lenguas.

En este apartado se aborda la conformación de los equipos multidisciplinarios y las funciones que realizó cada integrante en el trabajo de campo. También se

presentan las herramientas y los programas informáticos utilizados, así como las etapas de los procesos de registro de los datos preliminares y el procesamiento de los datos primarios.

2.1. Equipo de investigación

Se conformaron 3 equipos de trabajo, como se muestra en el Cuadro 2, compuestos por sabios representantes de los pueblos ashaninka, nomatsigenga y yanesha, docentes en educación intercultural bilingüe, comunicadores sociales y lingüistas. Los sabios de cada pueblo brindaron sus conocimientos en cada evento documentado, los docentes se encargaron de la entrevista y traducción en cada una de las lenguas originarias, los comunicadores sociales registraron los datos preliminares y los lingüistas se dedicaron al proceso de creación de anotaciones lingüísticas. La interrelación entre los miembros de cada equipo demostró que el trabajo colaborativo es fundamental en los procesos de documentación de lenguas (Cf. Glenn, 2009).

Los jóvenes docentes formados en distintas dimensiones de la interculturalidad con conocimiento de las lenguas originarias y con nociones de análisis lingüístico se convirtieron en actores claves al momento de interrelacionar a los sabios con los comunicadores sociales y lingüistas. Actuaron como los denominados lingüistas comunitarios dentro de la documentación lingüística (Cf. Mosel, 2007). El acercamiento generacional de estos jóvenes a las herramientas modernas facilitó el entrenamiento en el uso de programas informáticos como ELAN.

2.2. Herramientas

Cuadro 2. Equipo del Proyecto de Documentación CILA-UNMSM

	ashaninka	nomatsigenga	yanesha
Sabios de los pueblos	Seferino Ruiz	Violeta Mincami	Rolando Araujo
Docentes EIB	Yaneth Cameli Katerin Rojas	Katerine Campos Samuel Benito	Izhar Dionisio Darío Lopez
Comunicadores sociales	Milton Lopez	Jesús Castillo	Sofía Álvarez
Lingüistas	Licett Ramos	Pamela Jiménez	Rosario Cosar Pier Súclupe

Fuente: Informe del proyecto de investigación N.º 140304041 FLCH-UNMSM

Los archivos de audio y video que conformaron parte de los datos preliminares se registraron por medio de 2 cámaras Canon modelos EOS 60D y EOS Rebel T3i, con objetivos de 18 mm-55 mm, 50 mm y 28 mm-135 mm, una grabadora TASCAM DR-40 y un micrófono unidireccional HTDZ HT-81. Una de las cámaras registraba los primeros planos y la otra, los detalles en los ambientes que los mismos sabios indicaron como se puede observar en la Imagen 1. De esta manera, los datos obtenidos fueron de alta calidad de acuerdo con lo que sostiene Austin (2007) cuando dice que "un buen corpus documental debe incluir materiales en video y/o audio, de preferencia grabados en buenas condiciones y en locaciones auténticas" (p. 114).

Por otra parte, para obtener los datos primarios a través del procesamiento de los datos iniciales, se utilizaron computadoras y laptops a las que se les instaló los programas Arbil 2.5, ELAN 4.6.1, Adobe Premiere Pro y Canon Utilities Digital Photo Professional. En la Imagen 2 se muestra a la docente bilingüe Izhar Dionisio en el uso del programa ELAN cuando se encontrada transcribiendo en yanesha.

Imagen 1. Registro de audio y video para obtener los datos iniciales

Fuente: Archivo fotográfico del proyecto de investigación N.º 140304041 FLCH-UNMSM

2.3. Registro de datos preliminares

Este proceso estuvo bajo la dirección de los comunicadores sociales en coordinación con los sabios, docentes y lingüistas. Se grabaron principalmente 3 eventos comunicativos con una duración aproximada de 5 horas en total. Luego, los videos fueron editados en el programa Adobe Premiere y convertidos al formato mpg2 y los audios se mantuvieron en su formato original wav.

Los resultados se consignaron en 3 sesiones. La primera sesión, para el caso del ashaninka, se denominó *shipokantsi* y trató sobre un evento curativo tradicional que consiste en diagnosticar la dolencia y realizar la cura a través del vapor que emerge de plantas medicinales al contacto con piedras calientes. Para el caso del nomatsigenga, la sesión fue llamada *kitsagarentsi* y mostró la confección de una cushma de algodón acompañada de un relato y un canto. La última sesión, para el caso yanesha, se denominó *tamell* y registró la construcción de una trampa y el diálogo sostenido entre abuelo y nieto como parte del aprendizaje de esta actividad. La Imagen 1 captura un momento del registro de este último evento comunicativo.

2.4. Procesamiento de datos primarios

El proceso de anotación estuvo a cargo de los lingüistas y los docentes hablantes de cada una de las lenguas documentadas. Las anotaciones se realizaron a través del programa ELAN versión 4.6.1 (Sloetjes y Wittenburg, 2008). Dicho programa permitió etiquetar, transcribir y traducir la información lingüística de cada uno de los eventos registrados. El etiquetado de la información consistió en atribuir un nombre a cada línea. En este caso, se crearon tres líneas: una de transcripción (tx@), una de traducción (ft@) y otra de comentarios (com@). En la siguiente imagen, se observa el proceso de anotaciones para el registro de la sesión sesión *tamell*.

Para el proceso de la creación de anotaciones, se consideró como criterio de segmentación el límite entre las pausas indicadas por el oscilograma. Asimismo, se optó por utilizar una transcripción grafémica basada en los alfabetos normalizados por el Ministerio de Educación del Perú, una traducción libre propuesta por los docentes bilingües y comentarios que indicaron el uso de muletillas, repeticiones, descripción de glosas, contextos lingüísticos, entre otros datos.

Foto 1. Procesando datos primarios con anotaciones en Elan

Fuente: Archivo fotográfico del proyecto de investigación N.º 140304041 FLCH-UNMSM

3. Resultados y discusión

En este apartado se muestran ejemplos de algunos análisis lingüísticos que se pueden realizar a partir de los datos primarios obtenidos en los procesos de documentación de lenguas. En este caso, se presentan ejemplos para el análisis léxico y gramatical. En el análisis léxico, las entradas refieren principalmente a cada evento comunicativo como plantas curativas para la sesión *shipokantsi*, productos del tejido para la sesión *katsigarentsi* y especie de animales para sesión *tamell*. En el análisis gramatical, los datos primarios nos permiten realizar análisis a nivel de la frase verbal para el caso del ashaninka, a nivel de la frase nominal para el caso del yanesha y a nivel discursivo para el caso del nomatsigenga.

3.1. Datos primarios para la investigación léxica

De acuerdo con las sesiones analizadas, se puede obtener léxico especializado según el evento comunicativo registrado. En todos los casos, se muestra la entrada, la categoría léxica, la definición y, al parecer lo más interesante, los ejemplos en uso con traducción libre tomados del proceso de documentación lingüística. Para el caso de la sesión *shipokantsi*, las entradas mostradas refieren principalmente a plantas curativas; para el caso de la sesión *katsigarentsi*, las entradas se refieren a los productos del proceso de tejido y, para el caso de la sesión *tamell*,

Plantas curativas en la sesión shipokantsi

Iroshiki. s. Planta curativa cuyas propiedades curan el prolapso, la caída de cabello y estimulan el apetito durante la menstruación. ***Iroshiki*** *kametsatatsi aisati. '**Iroshiki** es bueno también'.*

Kajiroipini. s. Hierba curativa de comején contra, principalmente, desmayos y mareos. ***Kajiroipini*** *keme kajiro.* ***Hierba de comején*** *tiene olor de comején.*

Marankishi. s. Planta curativa denominada hoja de culebra. *Akantiro* ***marankishi****, ikepiyaneti yobanti. 'Se llama* ***hoja de culebra****, es venenosa y mata'.*

Pantiroshi. s. Planta curativa de flor grande de color violeta que contrarresta la esterilidad. *Eroka ikanteriro* ***pantiroshi****. 'Esto se llama* ***pantiroshi****'.*

Shibitsa. s. Bejuco curativo que sirve para combatir la gonorrea y la inflamación de la próstata. *Meka oka* ***shibitsa*** *aka tekatsi añero. 'Ahora este* ***bejuco curativo*** *aquí es escaso de encontrar'.*

Tonkitsiki. s. Hierba curativa contra el mal de los huesos. *Jeroka pashini atsipatajerori: tonkitsiki. 'Aquí está el otro con lo que le vamos a juntar:* ***hierba para los huesos****'.*

Productos del tejido en la sesión *kitsagarentsi*

Chogompërontsi. s. Tejido utilizado para cargar a bebés. *Ora nigotasantë antagaisati kitsagarentsi, tsagi aige* ***chogompërontsi****. ´Sí, lo que aprendí es hilar de todo, por ejemplo, la cushma, tsarato y* ***cargador de bebé*** *´.*

Kitsagarentsi. s. Cushma de hilo. *Aisati ora* ***kitsagarentsi****. ´También como la* ***cushma de hilo*** *´.*

Manchagintsi. s. Cushma de tela. *Pairi ira kara pintirori oka **manchagintsiga** kitsagarentsika. ´Cuando haces la **cushma**, ¿con quién lo haces?´*

Tsagi. s. Tsarato (bolso). *Ora nigotasantë antagaisati kitsagarentsi, **tsagi** aige chogompërontsi. ´Sí, lo que aprendí es hilar de todo, por ejemplo, la cushma, **tsarato** y cargador de bebé´.*

Especie de animales en la sesión *Tamell*

Ashath. s. Especie de paloma pequeña. *Añetholla yetamellpa' añethollachña (añethollachña) rrema', eteth ñerram **ashath**... 'Si lo construimos pequeño atrapa, únicamente, animales pequeños, como **paloma**'.*

Posholl. s. Especie de ardilla colorada. *Oñem **posholl**pa' allchaxho... 'Por ahí pasa la **ardilla**…'.*

Rroth. s. *Especie de ave grande como la pava. Terra eñall allampaño **rroth**pa', ochkaxho' eso'the mothanmats ñamha. 'Si hay cerca la **ardilla**, también se le atrapa'.*

Somh. s. Especie de perdiz de tamaño mediano, color negro. *Allpa' eñall **somh** 'Ahí hay **perdiz**'.*

las entradas pertenecen especies de animales en la caza.

3.2. Datos primarios para la investigación gramatical

Los datos primarios obtenidos en la documentación lingüística pueden ser utilizados también para realizar análisis gramaticales. En los siguientes ejemplos se muestran posibilidades de análisis a nivel de la morfosintaxis o a nivel discursivo. Al igual que el caso anterior, los ejemplos que se muestran están basados en el uso de las lenguas originarias en los procesos de documentación a los cuales se les ha añadido la descripción gramatical.

La frase verbal del ashaninka permite que se adjunte una serie de morfemas a su

raíz, aproximadamente, 40 de acuerdo con Mihas (2010). En los siguientes casos, se muestran ejemplos de los morfemas -rika, -pero, -ako y -ant que indican

Morfema -rika (condicional/COND): cuando se adhiere a un verbo, este deviene en una condición subordinada.

Timatsirika oshitsane, ari añakero.
Timatsi- **rika** oshitsane, ari a- ñ - ak - e - ro.
V - **COND**. N. Det- 1p.pl. V - PERF- T- OD
'**Si** tiene parásitos, aquí lo veremos'.

Morfema -pero (intensificador/INTENS): intensifica la acción verbal.

Pimoijakaperote
Pi- moijaka-**pero** - t - e
2 PS. V – **INTENS**- epéntesis - TNF
'Le haces hervir **bien**'

Morfema -ako (dativo/DAT): indica 'con', 'de', 'acerca de', 'sobre'

Pikobakotanaro.
Pi - kob - **ako** - t - an - a - ro
2 P.S - V - **DAT**. - epéntesis – DIR – T - OD
'Me estás pidiendo que te enseñe [**sobre algo**]'.

Pipokashitakotanari meka aka nonampiki.
Pi - pok - ashi - t - **ako** - t - an - a - ri
2 P.S - V PROP - epéntesis – **DAT** - epéntesis – DIR– T- OD
'Vienes a/**para** aprender [**acerca de algo**] aquí (aquí) en mi pueblo'.

Morfema -ant (resultativo/RESULT): indica una acción que presenta justificación.

Piyotantyari.
Pi - yot - **anty** - a - ri
2 P.S - V - **RESUL**. - T - SUB.
'Para que sepas'.

Otimantanajeari otomí.
O - tim - **ant** - an- aj- e - a - ri otomí.
3 P.S. F - V - **RESUL**.- DIR. – PERF. - T – epéntesis- OD.
'Para tener otro hijo'.

el condicional, intensificador, dativo y resultativo como lo señala Payne (1989). Para el caso del yanesha, de acuerdo con Duff-Tripp (2008), el plural en los sustantivos se marca con los sufijos -(V)*xhno/xhno* 'varios tipos de' y -esha' 'un grupo de'. Los términos que se pluralizan con *-esha'* son 1) los que refieren parentesco, 2) los objetos pequeños (siguiendo al diminutivo -(V)*tholl*), 3) los grupos de personas de la misma categoría (siguiendo a los sufijos *-Vn/-n/-nan* 'gente') y 4) los animales o cosas consideradas personas. En el corpus analizado, encontramos 7 casos de uso del sufijo *-xhno* y 20 casos de uso del sufijo *–esha'*. De este último, encontramos que *–esha'* pluraliza a tres de los cuatro tipos de términos

Tabla N.° 1. Aparición de -esha' según términos pluralizados

		Parentesco	Diminutivo	'Gente'	Ant. Pers.
Rolando	10	4	4	2	-
Dario	10	4	-	6	-

mencionados por Duff-Tripp (Ver Tabla N.° 1).
En esta distribución, observamos que Rolando y Dario emplean la misma cantidad de veces términos de parentesco pluralizados con *–esha'* (p. ej. *yaneshaypa'* 'los *yaneshas*') y en ninguna ocasión usan este sufijo para referir objetos que anteriormente fueron personas. La diferencia entre ambos la encontramos en que Rolando emplea –esha' cuatro veces en palabras que tienen el diminutivo -tholl (p. ej. ápexhothollesha '(las) soguillas'), mientras que Dario no lo hace en ninguna ocasión. Asimismo, notamos otra diferencia en que Dario usa más veces que Rolando el sufijo *–esha'* (6-2) para referir 'gente' (p. ej. *cheshachoresha'* 'los jóvenes').

Estas dos diferencias se correlacionan con el tipo de discurso: en el primer caso, Rolando usa más *–esha'* asociado al diminutivo *–tholl* para referir a los objetos y animales vinculados a la caza con el *tamell*, quizá para dar énfasis y así hacer comprensible dichos elementos. En el segundo caso, Dario emplea *–esha'* para referir 'gente' cada vez que realiza una pregunta en las que el sujeto es genérico (p. ej. *¿Eseshatña shomemhtetsa?, ¿ñato enkanesha' ñato koyanesha' ñato cheshachoresha'?* '¿Quiénes cazan, los varones, las mujeres o los jóvenes?').

A nivel discursivo, las siguientes partículas del nomatsigenga aike, aro y kara indican secuencia cronológica, secuencia lógica, medio o propósito o relaciones de ambiente. Los ejemplos fueron obtenidos de los datos primarios de la sesión

Aike: De acuerdo con Shaver (2008), indica secuencia espacial o temporal entre eventos independientes.

*irasi sërari **aike** tsinane nega okanta*
´ese la del varón y (**después**) de las mujeres cómo es´

Aro: ´así que´, ´ya que´

*anangerota nega **aro** omarintanëmarori naro yamë nobairo*
´se hace **de esta manera así** se inicia a elaborar ahora, mi nombre es´

Kara: ´a´

*nomarintëmaroti yamë **kara** okanta ganiri okantagani nega*
´ahora voy **a** iniciar´

*paniroma aige onetsëgapëna nasintojegira omanigiapë **kara** nakantirora*
´Lo hago sola, pero sí vienen mis hijas **a** mirar cuando yo estoy haciendo´

Kara: ´allá´, ´donde´

ogontëinarora iroro irorokeraita komantaiganakerine **kara** *ogomeantëri iroro negaraka onenake ojage*

´para que ellos puedan aprender y puedan transmitir lo que también ella va a aprender **donde** quiera que vaya´

kitsagarentsi.}

4. Conclusiones

Las metodologías de trabajo de campo y el desarrollo de herramientas informáticas en la documentación lingüística nos ofrecen una línea metodológica interesante que puede extenderse a toda la práctica disciplinar en el campo de la lingüística. Si llegáramos a establecer una relación entre los objetivos de los estudios lingüísticos propios con los procesos de documentación de lenguas, la labor en nuestra disciplina sería bastante redituable, al menos por dos motivos. Por un lado, las metodologías de investigación mejorarían el recojo y evidencia

de nuestros corpus de datos, y, por otro, estaríamos contribuyendo a preservar las lenguas o variedades de esta que estén o no necesariamente amenazadas o en peligro de extinción.

En los procesos de documentación lingüística, se pueden resaltar dos aspectos metodológicos que se pueden extender a toda la disciplina en general. Uno de ellos se relaciona con el trabajo colaborativo entre distintos actores académicos y no académicos que permite que cada experto aporte al cumplimiento de un solo objetivo y el otro se relaciona con el uso de tecnologías que van desde herramientas para recoger datos hasta programas informáticos para procesarlos. Estos dos aspectos permiten mejorar el recojo y evidencia de los corpus utilizados en los análisis lingüísticos.

En este trabajo en particular, se ha logrado contar con evidencia audiovisual de tres eventos comunicativos de las lenguas ashaninka, nomatsigenga y yanesha que procesados con programas informáticos como ELAN nos han permitido analizar con más detalle el aspecto léxico y gramatical de estas tres lenguas. Es más los datos obtenidos en estos procesos de documentación no solo proveen información interesante para el análisis lingüístico, sino su utilidad también se extiende hacia otras áreas como la medicina, la biología, la antropología, entre otras.

Notas

[1] Los proyectos consignados en el Archivo ELAR (The Endangered Languages Archive at SOAS, London) pueden accederse a través de la siguiente dirección web: http://elar.soas.ac.uk/

[2] Los proyectos consignados en el Archivo DOBES (Documentation of Endangered Languages) pueden accederse a través de la siguiente dirección web: http://dobes.mpi.nl/

[3] Los proyectos consignados en el Repositorio Digital de Lenguas Peruanas http://repositorio.pucp.edu.pe/index/handle/123456789/15344

Referencias bibliográficas

Austin, P. (2007). Los datos y la documentación lingüística. En J. Haviland y J. Flores (Ed.), *Bases de la documentación lingüística*. (pp. 111-140) México, D.F.: Instituto Nacional de Lenguas Indígenas.

Dimmendaal, G. J. (2010). *Language description and «the new paradigm»: What linguists may learn from ethnocinematographers*. http://scholarspace.manoa.hawaii.edu/handle/10125/4475

Duff-Tripp, M. (2008). *Gramática del idioma Yanesha' (Amuesha)* (2a. ed.). En Wise, M. R. (Ed.) Serie Lingüística Peruana N.° 43. Lima: Ministerio de Educación, Instituto Lingüístico de Verano.

Glenn, A. (2009). *Five Dimensions of Collaboration: Toward a Critical Theory of Coordination and Interoperability in Language Documentation*. http://scholarspace.manoa.hawaii.edu/handle/10125/4437

Himmelmann, N. P. (2012). *Linguistic Data Types and the Interface between Language Documentation and Description*. http://scholarspace.manoa.hawaii.edu/handle/10125/4503

Himmelmann, N. P. (2007). La documentación lingüística: ¿qué es y para qué sirve? En J. Haviland y J. Flores (Ed.), *Bases de la documentación lingüística*. (pp. 11-47). México, D.F.: Instituto Nacional de Lenguas Indígenas.

Mosel, P. (2007). Los datos y la documentación lingüística. En J. Haviland & J. Flores (Ed.), *Bases de la documentación lingüística*. (pp. 91-110) , México, D.F.: Instituto Nacional de Lenguas Indígenas.

Shaver, H. (2008). *Diccionario Nomatsiguenga-castellano castellano-nomatsiguenga*. Lima, Perú: Instituto Lingüístico de Verano.

Sloetjes, H., y Wittenburg, P. (2008). *Annotation by category – ELAN and ISO DCR*. http://www.mpi.nl/publications/escidoc-60774/@@popup.